社交媒体数据在自然灾害应急管理中的应用研究综述

郭柯杰^{1,2}, 吴吉东^{1,2*}, 叶梦琪^{1,2}

(1. 北京师范大学地理科学学部环境演变与自然灾害教育部重点实验室, 北京 100875;

2. 应急管理部/教育部减灾与应急管理研究院, 北京 100875)

摘要: 社交媒体是灾害应急管理的一个新兴但尚未充分利用的大数据源。然而, 社交媒体数据在灾害应急管理中能发挥什么作用, 目前国内的研究较少, 而国外的研究亦较为分散, 相关应用受限。论文对国内外利用社交媒体数据进行灾害管理相关研究进行了综述, 归纳出社交媒体数据在灾害应急管理中应用的潜力、优势和问题。结果发现: ① 社交媒体数据含有丰富的灾害信息, 能够依靠态势感知和信息共享支持灾害预警、实时监测、损失和救助需求评估、协助快速应急响应以及预测可能的灾害风险等灾害管理工作; ② 社交媒体数据的优势在于能保证灾害信息的实时性和连续性, 进而提供致灾强度图、损失预测和舆情分析等结果服务于灾害应急管理; ③ 地理位置信息有限、专业语料库缺乏、信息噪声处理的复杂性等原因使社交媒体数据分析结果精度受限。如何发挥社交媒体数据社会感知的优势, 以弥补传统观测调查手段的不足, 更加科学应对灾害应急, 仍需要在社交媒体数据处理及多源数据融合分析技术方面取得突破。

关键词: 社交媒体; 自然灾害; 应急管理; 机器学习

全球每年有大量自然灾害发生, 自然灾害常常具有复杂性、不确定性及动态性等特点^[1-2]。严重的自然灾害往往会造成高昂经济损失和重大人员伤亡^[3-5]。重大自然灾害频发对区域的可持续发展来说, 不仅是沉重的压力, 也是严峻的挑战^[6-7]。要完全消除自然灾害的影响几乎是不可能的^[8], 但可以通过灾害管理来降低灾害损失^[9]。目前, 包括社交媒体、位置服务和机器学习等在内的新兴技术被证实可用于灾害管理^[5]。

随着互联网技术的发展和5G时代的来临, 大数据受到了越来越多的关注。李德仁^[10]总结出大数据具有速度快、体量大、真伪难辨、模态多样、价值大等特点, 并指出时空大数据的挖掘和研究将会是未来的热点。作为一种时空大数据, 社交媒体数

据具有实时性和位置服务的特点^[11], 已经成为灾害管理的研究对象之一^[12-13], 将社交媒体数据应用到灾害应急计划和危机管理中, 能有效提高应急管理的效率^[14]。

社交媒体数据能够弥补传统遥感、气象观测数据在解析致灾程度、灾情中的不足。遥感技术的优势在于实时监测灾害进程, 反馈图像和波段信息, 不足在于采集灾害数据的时间周期较长, 且遥感产品与受灾群众的交互性不高^[15]。气象观测数据的优势在于定点观测致灾强度信息, 进而采用数值天气预报模型以及大气模型等模拟致灾强度的空间分布, 不足在于模型模拟出的结果准确度不能保证, 且该结果仅代表水文气象方面的考虑^[16], 而实际受灾程度受地形、设防水平等多方面因素的影

收稿日期: 2019-11-04; 修订日期: 2020-04-05。

基金项目: 国家重点研发计划项目(2018YFC1508903); 国家自然科学基金项目(41571492)。[Foundation: National Key Research and Development Program of China, No. 2018YFC1508903; National Natural Science Foundation of China, No. 41571492.]

第一作者简介: 郭柯杰(1997—), 男, 硕士生, 主要从事自然灾害风险管理研究。E-mail: wukejiennu@163.com

*通信作者简介: 吴吉东(1981—), 男, 博士, 教授, 主要从事灾害经济学、自然灾害风险管理研究。E-mail: wujidong@bnu.edu.cn

引用格式: 郭柯杰, 吴吉东, 叶梦琪. 社交媒体数据在自然灾害应急管理中的应用研究综述 [J]. 地理科学进展, 2020, 39(8): 1412-1422. [Wu Kejie, Wu Jidong, Ye Mengqi. A review on the application of social media data in natural disaster emergency management. Progress in Geography, 2020, 39(8): 1412-1422.] DOI: 10.18306/dlkxjz.2020.08.014

响,因此模拟结果与实际受灾情况可能存在较大差异。志愿地理信息(Volunteered Geographic Information, VGI)的概念由 Goodchild 在 2007 年首次提出^[17],核心思想是“人人都是传感器”^[18],即每一个体都可以采集地理信息数据进行共享。VGI 通常是指普通用户以移动互联网为媒介自发协助完成空间地理信息数据的采集、处理、管理与维护^[19],其特点包括 Web 2.0^[20]、集体智慧^[21]与新地理^[22]。与 VGI 类似,相关研究证明社交媒体数据能反映灾情信息^[23-24],感知灾害信息^[25-32],其特殊表现在信息传递的及时性、数据的共享性及信息的感知性。

针对传统观测技术感知灾害信息的不足和社交媒体数据的优势,国内外学者开始探讨社交媒体数据在灾害应急管理中应用的技术方法。本文选取最新灾害社交媒体数据的研究成果,从灾害应急管理中的致灾、损失和救助需求出发,综述利用社交媒体数据提取灾害感知信息的方法与不足,讨论利用社交媒体数据进行灾害应急管理的优势和挑战,为后续相关研究提供参考。

1 社交媒体数据提取灾害信息的应用

1.1 基于社交媒体数据的灾害信息提取方法

社交媒体数据及信息提取研究为致灾强度、承灾体损失程度和救助需求评估提供了新的途径。

社交媒体数据提取致灾强度信息的方法有:①基于社交媒体数据得到致灾因子的强度空间分布情况,进而提取致灾强度信息^[33];②使用位置熵和马尔可夫转移矩阵结合时空维度,反映致灾因子动态演变^[34];③通过空间聚类识别受灾区^[35]乃至重灾区^[36]。特别地,对于地震灾害,可以利用空间增长模型快速估计震区烈度^[37],但该方法对原始数据要求较高。另外,滤波分类^[38]和时间序列与空间变化分析^[39]也能快速识别、定位和监测震区烈度,但其在精度上常劣于监测数据。因此,社交媒体数据常与监测数据结合,用于修正监测数据,得到更合理的地震烈度分布图^[40-41]。但在提取致灾强度信息的过程中,存在以下问题:①对文本信息的分类计数方法并不能应用到图像视频的处理中;②虽然保证了获取致灾强度信息的实时性,但没有考虑细化到街道及具体建筑物的受灾情况;③相关研究通过位置纠正提高了社交媒体数据的地理精度^[42],但由于缺乏官方的致灾强度图,使优化结果无法验证等。

在利用社交媒体数据提取因灾损失信息方面的主要方法有:①采用语义分析^[43]、语义分类^[44]、情感分析^[45]和主题标签^[46]等多维分析方法预测灾害损失,并用机器学习和反馈机制提高预测精度^[47];②对于损失细节信息的提取,信息检索技术^[48]和地理标记评分法都是可行的^[49];③航测照片和推文融合分析技术也能估计直接经济损失^[50]。但仍存在以下问题:①损失预测结果精度受限于数据质量;②在特定灾害中涉及更多的是间接经济损失,无法通过社交媒体数据进行定量估算^[50]等。

利用社交媒体数据提取灾害救助需求信息的方法主要有:①通过语义分析识别舆论走向,提供救灾参考^[27,51],并结合机器学习^[52]来获取需求信息^[53];②分析推文数量和空间的变化,为重灾区的救援提供保障^[54];③考虑人口类型^[55],分析公众情绪,寻找消极情绪的空间分布规律,进而分配救灾物资^[56]。主要存在的问题包括:①需求信息依赖于含有精确地理位置的社交媒体数据,大部分社交媒体数据无法满足要求;②社交媒体信息可能存在个人夸张成分,导致需求信息与实际受灾情况不匹配。

除上述灾害信息外,社交媒体数据还能提供交互信息。如 Tang 等^[57]对社交媒体数据在干旱管理中的功能进行了探索,通过分析干旱持续期间主流社交媒体平台的信息,发现用户在讨论与干旱相关的问题时,将可能的解决方案概念化,为政府相关部门所借用并实施;政府通过社交平台与受灾群众建立沟通,最终向公众提供更合理的行动方案。

社交媒体数据还有监测作用。如 Yin 等^[58]开发了利用社交媒体数据实时监控森林火灾的系统,通过计算含有“fire”等关键词的推文数量,监测火灾事件,并使用支持向量机方法对推文进行分类,进一步提取用户、提及次数和位置信息等。根据推文的相似度和发布时间进行分组,并在交互式地图上进行点位可视化,显示森林火灾发生的位置。但是此监控系统对推文的地理精度要求较高,使用含粗略位置的推文时可能会造成误差。

1.2 基于社交媒体数据提取灾害应急管理信息的案例

为弄清社交媒体数据能提供怎样的灾害信息,并且以怎样的方式呈现, Kryvasheyeu 等^[59]以 2012 年美国桑迪飓风为例,利用 Twitter 数据进行了飓风灾害损失评估,研究结果如图 1 所示:①对获取的 Twitter 数据进行时空分析可知,到飓风路径的距离越大,推文的强度越高;②为了确定 Twitter 活动能

注:修改自 Kryvasheyeu 等^[59]。

图1 利用Twitter数据对桑迪飓风灾害损失进行快速评估研究结果示意图

Fig.1 Schematic diagram of rapid loss assessment of hurricane Sandy using Twitter data

否定量反映灾害损失严重性,在基于情绪打分的基础上,作者还分别计算了Twitter活动—损失的相关性和情绪—损失的相关性,图1b的结果表明Twitter活动和灾害损失之间存在中等正相关关系,而Twitter情绪与损失之间不存在明显相关性;③结合美国联邦应急管理局(Federal Emergency Management Agency, FEMA)提供的家庭援助款项数据和保险索赔数据,图1a计算出损失值与Twitter活动强度之间存在显著的指数关系,证明人均Twitter活动与飓风造成的人均经济损失密切相关。

已有研究表明:虽然各研究角度不同,但技术方法类似,大多遵循图2所示的技术路线。主要技术流程包括:①数据源。多类社交媒体平台能够提供丰富的文本、位置、图片和视频信息。Twitter是研究最多的社交媒体平台^[60-63],由于Facebook和新浪微博的地理信息没有Twitter数据完善^[64],因此较少应用于灾害管理。Twitter只有低于2%的推文含有地理标签^[65],位置信息的来源包括2类:一是用户在Twitter上的注册地信息^[66],二是通过手机GPS获取的定位信息。在考虑位置信息进行应急管理时,二者都有劣势:注册地通常不对应于灾害发生时用户的位置;而GPS位置仅在GPS开启情况下才能获取^[67]。而灾害管理中致灾、救助、恢复重建等管理需求都需要高精度的位置信息,由于微博等数据的商业属性和个人信息隐私等问题,微博等社交媒体信息地理位置获取的问题可能是限制其在灾害应急管理中应用的瓶颈。②预处理。原始

社交媒体数据含有较多噪音,需要通过机器学习和分词技术删除与灾害主题无关的数据,实现数据降噪。另外,对于地理位置不准确的数据,需要通过标签匹配等方法纠正并提取地理信息。③数据分析。方法种类繁多,重点在于数据的合理利用,充分挖掘数据的潜在信息。如对于时空分析,仅仅分析时空变化是不够的,还需对发生此变化的原因和驱动力进行分析,结合灾害特点,进一步提取灾害应急管理信息。④应急管理应用。通过数据分析和可视化技术得到致灾强度、损失分布、舆情信息和需求信息等,服务于灾害应急管理。在应用中,决策者和受灾群众之间通过应急管理的4个阶段(防灾、备灾、响应和恢复重建)相互联系。社交媒体在其中的应用能有效提高应急管理的效率。

同时,已有研究对灾害应急管理中致灾、损失和救助需求等关键环节信息的提取存在以下问题:①应急管理信息提取精度不高,例如对致灾因子强度信息的提取首先通过对相关致灾关键词进行分级,进而区分致灾因子强度,由于对关键词—致灾强度分级存在主观性,且文本发布可能存在主观夸张的成分,导致致灾强度信息提取精度不高;②对于地理精度,目前还没有较好的方法实现地理位置的精确校正,这使得致灾因子的空间信息有偏差;③对于灾害应急管理信息的提取,已有研究呈现出方法多样但普适性不高的问题。对于某类灾害,通过特定的信息提取或机器学习的方法,能够获取丰富而有效的灾害信息,但鲜有文献能够综合各类灾

图2 社交媒体数据在灾害应急管理中应用的技术路线图

Fig.2 Technology roadmap of using social media data in disaster management

害的特点,提出一套适用于大部分灾害信息提取的具体方法。

2 社交媒体数据的优势、潜力和挑战

2.1 社交媒体数据的优势

通过以上文献综述可知,从社交媒体可获取的灾害应急管理优势信息包括:① 高实时性的致灾信息。以台风为例,台风期间的推文附带了风和雨的强度信息,能实时反映致灾因子强度,但仍无法替代站点观测数据,二者结合的气象数据产品可能会有较好的应用前景。② 实时灾情信息。目前的灾情信息一般是通过灾区考察以及上下级部门统计上报的渠道获得,时间滞后性较大。而通过社交媒体中传递的灾情信息,决策者可以对整体灾情有基本认识。另外,社交媒体灾情信息还能统计上报

提供参考,提高上报数据的准确度和可信度。③ 对接受灾者的救助需求信息。以地震为例,在地震发生后,需求信息通常需要实地考察和评估得到,且救助需求通常是一个群体的总体需求,并不能精确到个人。但社交媒体中的救助需求信息是个人的、具体的、对接受灾者的,再精确的统计与评估都比不上受灾者客观提出的需求。

2.2 社交媒体数据的应用潜力

社交媒体数据应用于灾害应急管理的2个重要目的是态势感知和信息共享^[68],其在灾害应急管理的4个阶段:防灾、备灾、响应和恢复重建中有应用潜力^[69]。

在防灾阶段,工程措施、保险再保险和减灾战略规划等受到较多关注。要实施以上措施与规划,可靠的灾害大数据是必要的。社交媒体数据为防灾阶段提供信息参考,其中的典型应用为手机信令

数据。通过收集灾害发生前后手机呼叫及等待的详细信息,能够估计区域内人口分布和社会经济状况,以进行风险评估。手机等移动数据可以帮助决策者了解群众行为,进而模拟灾害管理计划并进行评估^[70]。规划避难所也是防灾阶段的重要措施,Kusumo等^[71]分析了居民撤离情况,表明居民所需庇护所的位置与政府避难所的位置吻合度只有35.6%,该结果为避难所的选取提供了参考。

在备灾阶段,需要实施紧急物资储备、召集救灾队伍和通信后勤等预案措施来保障灾害响应的有效性,其中及时监测灾害进程是备灾阶段的必要措施。社交媒体数据能够监测灾害情况。如在台风来临时,人们通过发布社交媒体信息反应当地风雨情况。应急管理当局可以通过访问社交媒体平台,获得灾害基本情况,并建立起防灾备灾的情景意识^[72]。已有研究运用社交媒体进行降雨和洪水事件的监测,并通过网络地图应用于巴黎和伦敦等城市的备灾阶段^[73]。

在响应阶段,决策者需要通过指标统计对灾情有全局认识。目前国内主要的灾情统计来源于指标上报,但该方法存在时间滞后性。而社交媒体能支持统计指标的获取:①通过提取社交媒体中与致灾因子相关的文字和图片信息,可以快速得到致灾因子的分布和强度等指标。同样地,通过提取和统计与受灾严重程度相关的信息,可以初步反映社交媒体使用者附近的受灾情况。这些指标在整合上报数据时能作为信息参考。②损失的快速评估。如微博和Twitter都能被用于估计台风灾区的损坏程度^[74],FEMA还组织了公共和私人小组来分析灾害推文,进而找出需求点并进行援助,可见其在灾害响应中对于社交媒体数据的重视程度。此外,还有相关研究证明社交媒体数据和观测数据的结合可以提高前者的准确性,并提高后者的时间分辨率^[75]。综合数据在规划和应急救援方面非常有价值,这意味着越早提供该数据,死亡人数和财产损失将会降低得越多^[76]。③详细的需求信息。社交媒体数据来源于群众,其需求是精确到个人的。在进行资源分析和物资调度时,详细的需求规划能够提高救援效率。

在恢复重建阶段,各产业恢复情况、基础设施重建情况以及群众的心理变化等信息是决策者关注的。社交媒体能够显示恢复重建的进程和时空模式。如Yan等^[77]使用带有地理标签的社交照片,

通过可视化研究监测和评估2013年菲律宾地震和“海燕”台风之后旅游业的复苏情况。另外,通过机器学习方法能快速区分灾后重建阶段中的优先等级,有助于政府制定科学性和针对性的救助策略。此外,社交媒体还能反映和估计灾民实际位置及灾后讨论的主题等,帮助政府与民众适应灾后环境并加强灾害治理^[78]。

2.3 社交媒体数据的应用挑战

在灾害应急管理应用中,社交媒体数据的机遇与挑战并存。

(1)采集社交媒体数据不及时。目前基于关键词搜索是获取社交媒体数据的主要方法,考虑到灾害应急管理,该方法仍存在时间滞后。需要充分考虑推文发布的时间滞后性,综合微博内容中的时间信息,实现时间聚类或创建时间线^[79]。若能开发自动化社交媒体数据采集工具^[80-82],实时采集、分析和可视化社交媒体数据,对灾害应急管理来说是有利的。

(2)提取应急管理信息存在瓶颈。社交媒体数据包含大量文本、图片和视频信息,然而,如何从不同类型信息中提取与灾害应急管理相关的信息是一个巨大的挑战,例如缺乏从多媒体数据中提取灾害应急管理信息的专业语料库。灾害应急管理与舆情紧密联系,目前在舆情分析中,合适的语料库是保证舆情精度的前提。虽然腾讯、Google等商业公司已构建自己专门的语料库,如情感分析语料库,但对灾害应急管理来说,不同类型灾害的致灾程度、不同承灾体的破坏程度和救助需求感知信息表达形式都存在显著的差异。如何针对灾害应急管理需求,建立专业化的语料库以提取相应信息,是利用社交媒体数据进行应急管理的基础,而这种专业化语料库的缺乏是限制其信息提取精度的关键因素之一。

(3)分词的多义性问题。相比于英语,中文较为简练且内涵丰富,在处理推文时,必须考虑推文的多义及歧义等问题。目前主要的中文分词方法包括3类,分别为词典方法、统计方法和规则方法^[83],更综合的还有自动机分词算法^[84]、规则统计相结合的汉语分词算法^[85]和后缀数组无词典分词算法^[86]等。对于分词结果中多义性的处理,常用的方法为贝叶斯分类判别器,结合含有词义标注的大规模语料库资源,利用多义词在上下文中的特征概率给出歧义判别结果^[87]。尽管有丰富的中文分词方法,对于新

词的处理还是目前存在的难点之一。

(4) 灾害相关推文图片与视频的处理技术仍存在难度。对于表情图片的信息提取,情感分类器是一个有效方法,通过特征抽取,构建分类器以及性能优化的步骤能达到提取图片情绪的目的^[88]。在识别推文图像内容时,可通过支持向量机方法提取信息^[89]。但是对于灾区应急管理不同阶段拍摄的照片及视频的信息提取,如何更精确识别及提取灾情信息进行应急救援,不仅需要研究具体的图像识别及提取技术,还需要强大的数据(如PB级大数据)处理能力保障^[90],此方面的研究及灾害应急管理应用极少,尚存在技术难度。

(5) 单源社交媒体提取的应急管理信息有限,其精度往往难以达到科学管理决策的需求,如何利用文本、图片、视频以及常规观测等多源信息^[91],研发多源社交媒体信息融合分析技术^[92]以更加全面系统地提取灾害应急中的致灾、损失和救助需求等信息,仍将是未来研究的难点,也是未来的发展趋势。

3 结论

立足于社交媒体数据在感知灾害信息中的优势以及传统观测数据的不足,社交媒体数据在灾害应急管理方面受到越来越多的关注。本文选取最新的研究成果,综述了社交媒体数据在灾害应急管理中应用的研究进展。

社交媒体数据丰富的感知信息可以为灾害应急管理中致灾、灾情和救助信息的提取提供支持,具有传统观测调查数据无可比拟的大数据信息特征,一定程度上弥补了传统观测调查数据全面性、系统性的不足:① 社交媒体反映的致灾信息常常是实时的,可初步反映致灾因子的强度和分布,以及辅助修正地震、洪涝等灾害致灾强度分布图;② 社交媒体反映的及时灾情信息是灾害应急管理响应阶段的重要参考,已证实能在台风等灾害中修正和提高灾情统计上报结果的准确度;③ 态势感知信息可以在应急救助中发挥重要作用,决策者可以分析受灾群众的情绪和舆情走向,进而调整相关应急管理计划,提高应急管理效率。

基于态势感知和信息共享,社交媒体数据有望实现在灾害管理4个阶段:防灾、备灾、响应和恢复重建提供数据支持和信息参考。在防灾阶段,社交媒体数据能为减灾工程以及灾害保险再保险提供

数据支持,使防灾规划更具科学性;在备灾阶段,社交媒体数据能够实时监测灾害进程,反映灾害基本情况;在响应阶段,社交媒体数据能够提供灾情统计指标,包括致灾因子、损失评估和需求信息等,提高应急救援的效率;在恢复重建阶段,社交媒体数据能够反映群众总体的灾后恢复情况,为决策者监控重建进程提供参考。

社交媒体数据应用于灾害应急管理时,基本遵循以下技术流程:数据获取、预处理、数据分析和应急管理应用,关键技术在于:预处理和数据分析阶段。预处理阶段通过机器学习和分词技术等排除与灾害无关的信息,实现数据降噪,同时,为提取地理位置,还需要对有偏差的位置进行纠正,进而提高数据分析的精度。数据分析的难点在于地理位置存在偏差、专业语料库的缺乏以及图片视频的处理难度高等。

本文研究结果表明基于社交媒体数据和数据挖掘及机器学习方法,可以很好地提取灾害应急管理信息,而且传统观测手段很难像社交媒体数据那样能够获得灾害管理的社会感知信息。尽管社交媒体在灾害应急管理中已有较多研究,但其缺陷不容忽视。如何解决社交媒体数据地理位置模糊、灾害应急管理信息精度有限等问题,仍需要在社交媒体数据处理及多源数据融合分析技术方面取得突破。

参考文献(References)

- [1] Sun Q Y, Li X Y, Yu F. Designing an emergency continuity plan for a megacity government: A conceptual framework for coping with natural catastrophes [J]. *International Journal of Critical Infrastructure Protection*, 2016, 13: 28-35.
- [2] 王静爱, 史培军, 朱骊. 中国主要自然致灾因子的区域分异 [J]. *地理学报*, 1994, 49(1): 18-26. [Wang Jing'ai, Shi Peijun, Zhu Li. Regional variations of major natural disaster-causing factors in China. *Acta Geographica Sinica*, 1994, 49(1): 18-26.]
- [3] Altay N, Green III W G. OR/MS research in disaster operations management [J]. *European Journal of Operational Research*, 2006, 175(1): 475-493.
- [4] Galindo G, Batta R. Review of recent developments in OR/MS research in disaster operations management [J]. *European Journal of Operational Research*, 2013, 230(2): 201-211.
- [5] Akter S, Wamba S F. Big data and disaster management: A systematic review and agenda for future research [J]. *An-*

- nals of Operations Research, 2017, 283(2): 1-21.
- [6] 史培军, 孔锋, 叶谦, 等. 灾害风险科学发展与科技减灾 [J]. 地球科学进展, 2014, 29(11): 1205-1211. [Shi Peijun, Kong Feng, Ye Qian, et al. Disaster risk science development and disaster risk reduction using science and technology. *Advances in Earth Science*, 2014, 29(11): 1205-1211.]
- [7] 吴吉东, 何鑫, 王莱林, 等. 自然灾害损失分类及评估研究评述 [J]. 灾害学, 2018, 33(4): 157-163. [Wu Jidong, He Xin, Wang Cailin, et al. A review on classification and loss assessment of natural disasters. *Journal of Catastrophology*, 2018, 33(4): 157-163.]
- [8] Lin Moe T, Pathranarakul P. An integrated approach to natural disaster management: Public project management and its critical success factors [J]. *Disaster Prevention and Management*, 2006, 15(3): 396-413.
- [9] Eckhardt D, Leiras A, Thomé A M T. Systematic literature review of methodologies for assessing the costs of disasters [J]. *International Journal of Disaster Risk Reduction*, 2019, 33: 398-416.
- [10] 李德仁. 展望大数据时代的地球空间信息学 [J]. 测绘学报, 2016, 45(4): 379-384. [Li Deren. Towards geo-spatial information science in big data era. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(4): 379-384.]
- [11] 梁春阳, 林广发, 张明锋, 等. 社交媒体数据对反映台风灾害时空分布的有效性研究 [J]. 地球信息科学学报, 2018, 20(6): 807-816. [Liang Chunyang, Lin Guangfa, Zhang Mingfeng, et al. Assessing the effectiveness of social media data in mapping the distribution of typhoon disasters. *Journal of Geo-information Science*, 2018, 20(6): 807-816.]
- [12] Lindsay B R. Social media and disasters: Current uses, future options, and policy considerations [R/OL]. Congressional Research Service Report R41987. <https://fas.org/sgp/crs/homsec/R41987.pdf>. 2011.
- [13] Landwehr P M, Carley K M. Social media in disaster relief [M]// Wu W W. Data mining and knowledge discovery for big data. Berlin, Germany: Springer, 2014: 225-257.
- [14] Alexander D E. Social media in disaster risk reduction and crisis management [J]. *Science and Engineering Ethics*, 2014, 20(3): 717-733.
- [15] Joyce K E, Belliss S E, Samsonov S V, et al. A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters [J]. *Progress in Physical Geography*, 2009, 33(2): 183-207.
- [16] Jasper K, Gurtz J, Lang H. Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model [J]. *Journal of Hydrology*, 2002, 267(1-2): 40-52.
- [17] Goodchild M F. Citizens as sensors: The world of volunteered geography [J]. *GeoJournal*, 2007, 69(4): 211-221.
- [18] Goodchild M F. Citizens as voluntary sensors: Spatial data infrastructure in the world of Web 2.0 [J]. *International Journal of Spatial Data Infrastructures Research*, 2007, 2(2): 24-32.
- [19] 张红平, 顾学云, 熊萍, 等. 志愿者地理信息研究与应用初探 [J]. 地理信息世界, 2012, 10(4): 67-71. [Zhang Hongping, Gu Xueyun, Xiong Ping, et al. Research and application of volunteer geographic information. *Geomatics World*, 2012, 10(4): 67-71.]
- [20] Scharl A, Tochtermann K. The geospatial Web: How geobrowsers, social software and the Web 2.0 are shaping the network society [M]. New York, USA: Springer, 2007.
- [21] Smith J B. Collective intelligence in computer-based collaboration [M]. Washington D C, USA: American Psychological Association, 1996.
- [22] Turner A. Introduction to neogeography [M]. Sebastopol, USA: O'Reilly Media, 2006.
- [23] Hjorth L, Kim K Y. The mourning after: A case study of social media in the 3.11 earthquake disaster in Japan [J]. *Television & New Media*, 2011, 12(6): 552-559.
- [24] Houston J B, Hawthorne J, Perreault M F, et al. Social media and disasters: A functional framework for social media use in disaster planning, response, and research [J]. *Disasters*, 2015, 39(1): 1-22.
- [25] Starbird K, Palen L, Hughes A L, et al. Chatter on the red: What hazards threat reveals about the social life of microblogged information [C]// ACM. Proceedings of the 2010 ACM conference on computer supported cooperative work. New York, USA: Association for Computing Machinery, 2010.
- [26] Vieweg S, Hughes A L, Starbird K, et al. Microblogging during two natural hazards events: What twitter may contribute to situational awareness [C]// ACM. Proceedings of the SIGCHI conference on human factors in computing systems. New York, USA: Association for Computing Machinery, 2010.
- [27] Acar A, Muraki Y. Twitter for crisis communication: Lessons learned from Japan's tsunami disaster [J]. *International Journal of Web Based Communities*, 2011, 7(3): 392-402.
- [28] Bruns A, Burgess J, Crawford K, et al. #qldfloods and @QPSMedia: Crisis communication on Twitter in the 2011 South East Queensland floods [R/OL]. <https://eprints.qut.edu.au/48241/>. ARC Centre of Excellence for Creative Industries and Innovation. Brisbane, Australia:

- Queensland University of Technology, 2012.
- [29] Dugdale J, Van de Walle B, Koeppinghoff C. Social media and SMS in the Haiti earthquake [C]// ACM. Proceedings of the 21st international conference on world wide web. New York, USA: Association for Computing Machinery, 2012.
- [30] Starbird K, Muzny G, Palen L. Learning from the crowd: Collaborative filtering techniques for identifying on-the-ground Twitterers during mass disruptions [C]// ACM. Proceedings of 9th international conference on information systems for crisis response and management. Vancouver, Canada: Simon Fraser University, 2012.
- [31] Imran M, Elbassuoni S, Castillo C, et al. Extracting information nuggets from disaster-related messages in social media [C]// ACM. Proceedings of the tenth international conference on information systems for crisis response and management (ISCRAM'13). Baden-Baden, Germany, 2013: 12-15.
- [32] Neubaum G, Rösner L, Rosenthal-von der Pütten A M, et al. Psychosocial functions of social media usage in a disaster situation: A multi- methodological approach [J]. Computers in Human Behavior, 2014, 34: 28-38.
- [33] Mersham G. Social media and public information management: The September 2009 tsunami threat to New Zealand [J]. Media International Australia, 2010, 137(1): 130-143.
- [34] Wang Z, Ye X. Space, time, and situational awareness in natural hazards: A case study of Hurricane Sandy with social media data [J]. Cartography and Geographic Information Science, 2019, 46(4): 334-346.
- [35] Bakillah M, Li R Y, Liang S H L. Geo-located community detection in Twitter with enhanced fast-greedy optimization of modularity: The case study of typhoon Haiyan [J]. International Journal of Geographical Information Science, 2015, 29(2): 258-279.
- [36] Arthur R, Boulton C A, Shotton H, et al. Social sensing of floods in the UK [J]. PLoS One, 2018, 13(1): e0189327. doi: 10.1371/journal.pone.0189327.
- [37] Wang Y, Ruan S, Wang T, et al. Rapid estimation of an earthquake impact area using a spatial logistic growth model based on social media data [J]. International Journal of Digital Earth, 2019, 12(11): 1265-1284.
- [38] Sakaki T, Okazaki M, Matsuo Y. Earthquake shakes Twitter users: Real-time event detection by social sensors [C]// ACM. Proceedings of the 19th international conference on world wide web. New York, USA: Association for Computing Machinery, 2010.
- [39] Crooks A, Croitoru A, Stefanidis A, et al. Earthquake: Twitter as a distributed sensor system [J]. Transactions in GIS, 2013, 17(1): 124-147.
- [40] 孙振凯. PAGER: 地震影响的快速评估(USGS 情况说明书第 2010-3036 号) [J]. 国际地震动态, 2013(1): 4-7. [Sun Zhenkai. PAGER: Rapid assessment of the impact of earthquakes (USGS fact sheet No. 2010-3036). Recent Developments in World Seismology, 2013(1): 4-7.]
- [41] Resch B, Usländer F, Havas C. Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment [J]. Cartography and Geographic Information Science, 2018, 45(4): 362-376.
- [42] Ogie R I, Forehead H. Investigating the accuracy of geo-referenced social media data for flood mapping: The Pet-aJakarta.org case study [C]// Hellingrath B, Gojmerac R. Proceedings of the 2017 4th International Conference on Information and Communication Technologies for Disaster Management (ICT- DM). New York, USA: IEEE, 2017: 30-35.
- [43] Yuan F, Liu R. Feasibility study of using crowdsourcing to identify critical affected areas for rapid damage assessment: Hurricane Matthew case study [J]. International Journal of Disaster Risk Reduction, 2018, 28: 758-767.
- [44] Kongthon A, Haruechaiyasak C, Pailai J, et al. The role of social media during a natural disaster: A case study of the 2011 Thai flood [J]. International Journal of Innovation and Technology Management, 2014, 11(3): 2227-2232.
- [45] Wu D, Cui Y. Disaster early warning and damage assessment analysis using social media data and geo-location information [J]. Decision Support Systems, 2018, 111: 48-59.
- [46] Fang J, Hu J, Shi X, et al. Assessing disaster impacts and response using social media data in China: A case study of 2016 Wuhan rainstorm [J]. International Journal of Disaster Risk Reduction, 2019, 34: 275-282.
- [47] Enenkel M, Saenz S M, Dookie D S, et al. Social media data analysis and feedback for advanced disaster risk management [J/OL]. Computing Research Repository, 2018: 1802.02631. <https://arxiv.org/abs/1802.02631>.
- [48] Priya S, Bhanu M, Dandapat S K, et al. Characterizing infrastructure damage after earthquake: a split-query based IR approach [C]// ACM. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). New York, USA: Association for Computing Machinery, 2018: 202-209.
- [49] Li X, Caragea D, Zhang H, et al. Localizing and quantifying damage in social media images [C]// ACM. 2018

- IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). New York, USA: Association for Computing Machinery, 2018: 194-201.
- [50] Schnebele E, Cervone G, Waters N. Road assessment after flood events using non-authoritative data [J]. *Natural Hazards and Earth System Sciences*, 2014, 14(4): 1007-1015.
- [51] Vo Bao-Khanh H, Collier N. Twitter emotion analysis in earthquake situations [J]. *International Journal of Computational Linguistics and Applications*, 2013, 4(1): 159-173.
- [52] Sadri A M, Hasan S, Ukkusuri S V, et al. Crisis communication patterns in social media during Hurricane Sandy [J]. *Transportation Research Record*, 2018, 2672(1): 125-137.
- [53] Varga I, Sano M, Torisawa K, et al. Aid is out there: Looking for help from tweets during a large scale disaster [C]// Schuetze H, Fung P, Poesio M. *Proceedings of the 51st annual meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Sofia, Bulgaria, 2013: 1619-1629.
- [54] Martín Y, Li Z, Cutter S L. Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew [J]. *PLoS One*, 2017, 12(7): e0181701. doi: 10.1371/journal.pone.0181701.
- [55] Kumar D, Ukkusuri S V. Utilizing geo-tagged tweets to understand evacuation dynamics during emergencies: A case study of Hurricane Sandy [C]// Champin P, Gandon F, Médini L. *Companion Proceedings of the the web conference 2018*. Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018: 1613-1620.
- [56] Gruebner O, Lowe S, Sykora M, et al. Spatio-temporal distribution of negative emotions in New York City after a natural disaster as seen in social media [J]. *International Journal of Environmental Research and Public Health*, 2018, 15(10): 2275. doi: 10.3390/ijerph15102275.
- [57] Tang Z, Zhang L, Xu F, et al. Examining the role of social media in California's drought risk management in 2014 [J]. *Natural Hazards*, 2015, 79(1): 171-193.
- [58] Yin J, Lampert A, Cameron M, et al. Using social media to enhance emergency situation awareness [J]. *IEEE Intelligent Systems*, 2012, 27(6): 52-59.
- [59] Kryvasheyev Y, Chen H, Obradovich N, et al. Rapid assessment of disaster damage using social media activity [J]. *Science Advances*, 2016, 2(3): e1500779. doi: 10.1126/sciadv.1500779.
- [60] Earle P S, Bowden D C, Guy M. Twitter earthquake detection: Earthquake monitoring in a social world [J]. *Annals of Geophysics*, 2012, 54(6): 708-715.
- [61] Morss R E, Demuth J L, Lazrus H, et al. Hazardous weather prediction and communication in the modern information environment [J]. *Bulletin of the American Meteorological Society*, 2017, 98(12): 2653-2674.
- [62] Zou L, Lam N S N, Cai H, et al. Mining Twitter data for improved understanding of disaster resilience [J]. *Annals of the American Association of Geographers*, 2018, 108(5): 1422-1441.
- [63] Barker J L P, Macleod C J A. Development of a national-scale real-time Twitter data mining pipeline for social geodata on the potential impacts of flooding on communities [J]. *Environmental Modelling & Software*, 2019, 115: 213-227.
- [64] Liu F, Xu D. Social roles and consequences in using social media in disasters: A structural perspective [J]. *Information Systems Frontiers*, 2018, 20(4): 693-711.
- [65] Laylavi F, Rajabifard A, Kalantari M. A multi-element approach to location inference of twitter: A case for emergency response [J]. *ISPRS International Journal of Geo-Information*, 2016, 5(5): 56. doi: 10.3390/ijgi5050056.
- [66] Gao H, Barbier G, Goolsby R. Harnessing the crowdsourcing power of social media for disaster relief [J]. *IEEE Intelligent Systems*, 2011, 26(3): 10-14.
- [67] 曹劲舟, 武红宇. 基于微博位置签到数据的POI更新方法 [J]. *地理空间信息*, 2013, 11(2): 15-18. [Cao Jingzhou, Wu Hongyu. POI update method based on weibo location check-in data. *Geospatial Information*, 2013, 11(2): 15-18.]
- [68] Sakaki T, Matsuo Y, Kurihara S, et al. The possibility of social media analysis for disaster management [C]// IEEE. *2013 IEEE Region 10 Humanitarian Technology Conference*. New York, USA: IEEE, 2013: 238-243.
- [69] Yu M, Yang C, Li Y. Big data in natural disaster management: A review [J]. *Geosciences*, 2018, 8(5): 165. doi: 10.3390/geosciences8050165.
- [70] Ghurye J, Krings G, Frias-Martinez V. A framework to model human behavior at large scale during natural disasters [C]// IEEE. *2016 17th IEEE International Conference on Mobile Data Management (MDM)*. New York, USA: IEEE, 2016: 18-27.
- [71] Kusumo A N L, Reckien D, Verplanke J. Utilising volunteered geographic information to assess resident's flood evacuation shelters. Case study: Jakarta [J]. *Applied Geography*, 2017, 88: 174-185.
- [72] Chan J C. The role of social media in crisis preparedness,

- response and recovery [R/OL]. <http://www.oecd.org/governance/risk/The%20role%20of%20Social%20media%20in%20crisis%20preparedness,%20response%20and%20recovery.pdf>. OECD Report, 2012: 1-16.
- [73] Feng Y, Sester M. Extraction of pluvial flood relevant volunteered geographic information (VGI) by deep learning from user generated texts and photos [J]. *ISPRS International Journal of Geo-Information*, 2018, 7(2): 39. doi: 10.3390/ijgi7020039.
- [74] Virtual Social Media Working Group, DHS First Responders Group. Lessons learned: Social media and hurricane sandy [R/OL]. Homeland Security: Science and Technology. <https://www.hsdl.org/?view&did=739332>. Washington D C, USA: US Department of Homeland Security, 2013.
- [75] Cervone G, Sava E, Huang Q, et al. Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study [J]. *International Journal of Remote Sensing*, 2016, 37(1): 100-124.
- [76] Liou Y A, Kar S K, Chang L. Use of high-resolution FORMOSAT-2 satellite images for post-earthquake disaster assessment: A study following the 12 May 2008 Wenchuan Earthquake [J]. *International Journal of Remote Sensing*, 2010, 31(13): 3355-3368.
- [77] Yan Y, Eckle M, Kuo C L, et al. Monitoring and assessing post-disaster tourism recovery using geotagged social media data [J]. *ISPRS International Journal of Geo-Information*, 2017, 6(5): 144. doi: 10.3390/ijgi6050144.
- [78] 周利敏, 刘和健. 人工智能时代的社交媒体与灾害治理: 兼论国际案例 [J]. 理论探讨, 2019(6): 175-181. [Zhou Limin, Liu Hejian. On social media and disaster management in the age of artificial intelligence and international cases. Theoretical Discussion, 2019(6): 175-181.]
- [79] 徐伟, 赵斌, 吉根林. 基于滑动窗口的微博时间线摘要算法 [J]. 数据采集与处理, 2017, 32(3): 523-532. [Xu Wei, Zhao Bin, Ji Genlin. A time line summary algorithm based on sliding window. Data Acquisition and Processing, 2017, 32(3): 523-532.]
- [80] Becker H, Naaman M, Gravano L. Beyond trending topics: Real-world event identification on twitter [C]// Adamic L A, Baeza-Yates R, Counts S. Proceedings of the fifth international conference on weblogs and social media. Barcelona, Spain, 2011: 438-441.
- [81] Huang Y, Dong H, Yesha Y, et al. A scalable system for community discovery in twitter during hurricane sandy [C]// IEEE. 2014 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. New York, USA: IEEE, 2014: 893-899.
- [82] Win S S M, Aung T N. Target oriented tweets monitoring system during natural disasters [C]// IEEE. 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS). New York, USA: IEEE, 2017: 143-148.
- [83] 孙铁利, 刘延吉. 中文分词技术的研究现状与困难 [J]. 信息技术, 2009, 33(7): 187-189. [Sun Tieli, Liu Yanji. Research status and difficulties of Chinese word segmentation technology. Information Technology, 2009, 33(7): 187-189.]
- [84] 吴建胜, 战学刚, 迟呈英. 一种基于自动机的分词方法 [J]. 计算机工程与应用, 2005, 41(8): 81-82. [Wu Jiansheng, Zhan Xuegang, Chi Chengying. A word segmentation method based on automata. Computer Engineering and Applications, 2005, 41(8): 81-82.]
- [85] 赵伟, 戴新宇, 尹存燕, 等. 一种规则与统计相结合的汉语分词方法 [J]. 计算机应用研究, 2004, 21(3): 23-25. [Zhao Wei, Dai Xinyu, Yin Cunyan, et al. A Chinese word segmentation method combining rules and statistics. Computer Application Research, 2004, 21(3): 23-25.]
- [86] 张长利, 赫枫龄, 左万利. 一种基于后缀数组的无词典分词方法 [J]. 吉林大学学报(理学版), 2004, 42(4): 548-553. [Zhang Changli, He Fengling, Zuo Wanli. A lexicographical word segmentation method based on suffix array. Journal of Jilin University (Science Edition), 2004, 42(4): 548-553.]
- [87] 刘商飞, 张志祥. 基于改进的 Bayes 判别法的中文多义词消歧 [J]. 计算机与数字工程, 2009, 37(10): 32-35. [Liu Shangfei, Zhang Zhixiang. Chinese polysemy disambiguation based on the improved Bayes discrimination method. Computer & Digital Engineering, 2009, 37(10): 32-35.]
- [88] 张珊, 于留宝, 胡长军. 基于表情图片与情感词的中文微博情感分析 [J]. 计算机科学, 2012, 39(S3): 146-148. [Zhang Shan, Yu Liubao, Hu Changjun. Emotion analysis of Chinese microblog based on emoticons and emoticons. Computer Science, 2012, 39(S3): 146-148.]
- [89] 陈永波. 微博图像文字内容识别与感知 [D]. 北京: 北京交通大学, 2014. [Chen Yongbo. Weibo image text content recognition and perception. Beijing, China: Beijing Jiaotong University, 2014.]
- [90] 关雪峰, 曾宇媚. 时空大数据背景下并行数据处理分析挖掘的进展及趋势 [J]. 地理科学进展, 2018, 37(10): 14-27. [Guan Xuefeng, Zeng Yumei. Development and trend of parallel data processing in the context of spatio-temporal big data. Progress in Geography, 2018, 37(10): 14-27.]

- [91] 杨喜平, 方志祥. 移动定位大数据视角下的人群移动模式及城市空间结构研究进展 [J]. 地理科学进展, 2018, 37(7): 18-27. [Yang Xiping, Fang Zhixiang. Research progress of crowd moving mode and urban spatial structure from the perspective of mobile positioning big data. *Progress in Geography*, 2018, 37(7): 18-27.]
- [92] Li Z, Wang C, Emrich C T, et al. A novel approach to leveraging social media for rapid flood mapping: A case study of the 2015 South Carolina floods [J]. *Cartography and Geographic Information Science*, 2018, 45(2): 97-110.

A review on the application of social media data in natural disaster emergency management

WU Kejie^{1,2}, WU Jidong^{1,2*}, YE Mengqi^{1,2}

(1. Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education of China, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; 2. Academy of Disaster Reduction and Emergency Management, Ministry of Emergency Management & Ministry of Education, Beijing 100875, China)

Abstract: Social media is a data source that is emerging but underutilized for disaster emergency management. There are few existing studies on the role of social media data in disaster emergency management in China, and international studies are also relatively scattered, which make relevant applications limited. This article reviewed the research on disaster management with social media data. We summarized the potential, advantages, and problems of applying social media data in disaster emergency management. The results show that: 1) Social media data are full of disaster information. They can support disaster early warning, real-time monitoring, and loss and rescue needs assessment, assist rapid emergency response, and predict possible disaster risks that relies on situational awareness and timely information. 2) The advantage of social media data is that it can ensure the timeliness and continuity of disaster information acquisition, which can provide disaster intensity maps, loss prediction, and public opinion analysis for disaster emergency management. 3) Some problems limit the accuracy of social media data analysis results, including limited geographical location information, lack of professional corpus, and complexity of information noise processing, among others. In order to make full use of the advantages of social media data to overcome the shortcomings of traditional observation and investigation methods and respond to disasters more scientifically, breakthroughs in social media data processing and multi-source data fusion analysis technology are needed.

Keywords: social media; natural disaster; emergency management; machine learning